

## Creating Metadata Records

### Introduction

Information specialists concern themselves with the proper input, structure, storage, and most importantly, retrieval of information. Stored information is worthless unless individuals with need have the ability to retrieve the information in an understandable format. The recent proliferation of information from electronic sources presents information specialists with new challenges, one of which is the arrangement of information in a machine-readable format to allow users great latitude in search and retrieval. Metadata, structured data about data, is the term given to the information collected about the electronic sources. Two systems developed to aid in the standardization of metadata are Electronic Resource Citation (ERC) and Dublin Core (DC).

### Electronic Resource Citation (ERC)

Electronic Resource Citation (ERC), developed by John A. Kunze, was created in response to Mr. Kunze's concern with maintaining direct human contact in the creation of metadata. In its simplest form, ERC requires only four basic elements to create a story about the resource. Those four elements: who, what, when, and where, are then considered a segment. ERC records may include many segments, all providing further insight into the information entity being described. The foremost ERC segment is called the anchoring story which is usually the main basic reference to the entity in terms of the user or audience determined to be of most import. Five simple examples of ERC follow:

1. Understanding Metadata. National Standards Information Organization. 2004

```
erc:  
who: National Standards Information Organization  
what: Understanding Metadata  
when: 2004  
where: http://www.niso.org/standards/resource/understandingmetadata.pdf
```

2. Brave New FRBR World. Patrick Le Bœuf. 2003

```
erc:  
who: Le Bœuf, Patrick  
what: Brave New FRBR World  
when: 2003  
where: http://www.ddb.de/news/pdf/papers\_leboeuf.pdf
```

3. The Digital Object Identifier System. [Website]

```
erc:  
who: The International DOI Foundation (IDF)
```

what: The Digital Object Identifier System  
 when: 20041123  
 where: <http://www.doi.org/index.html>

4. Trusted Digital Repositories: Attributes and Responsibilities. RLG-OCLC. 2002.

erc:  
 who: RLG  
 what: Trusted Digital Repositories: Attributes and Responsibilities  
 when: 200205  
 where: <http://www.rlg.org/longterm/repositories.pdf>

5. An Introduction to the Resource Description Framework. Eric Miller. 1998

erc:  
 who: Miller, Eric  
 what: An Introduction to the Resource Description Framework  
 when: 199805  
 where: <http://www.dlib.org/dlib/may98/miller/05miller/html>

### Dublin Core Metadata (DC)

In comparison to the four elements specified in ERC, Dublin Core has 15 elements in its simple set. Each of the 15 elements is optional and may be repeated as often as necessary. Originally established to describe networked document type resources, Dublin Core is applicable to various types of resources. Five examples of Dublin Core Metadata, corresponding to the five examples provided in ERC, follow:

1. Understanding Metadata. National Standards Information Organization. 2004

```
<META NAME="DC.Title" LANG="en" CONTENT="Understanding Metadata">
<META NAME="DC.Creator" LANG="en" CONTENT="National Information Standards
  Organization">
<META NAME="DC.Subject" LANG="en" CONTENT="metadata">
<META NAME="DC.Description" LANG="en" CONTENT="Table of contents; What is
  Metadata?; What Does Metadata Do?; Structuring Metadata; Metadata
  Schemes and Element Sets; Dublin Core; TEI and METS; MODS; EAD and LOM;
  <indec>, ONIX, CDWA, and VRA; MPEG; FGDC and DDI; Creating Metadata;
  Interoperability and Exchange of Metadata; Future Directions; More
  Information on Metadata; Glossary">
<META NAME="DC.Publisher" LANG="en" CONTENT="NISO Press">
<META NAME="DC.Contributor" LANG="en" CONTENT="Guenther, Rebecca">
<META NAME="DC.Contributor" LANG="en" CONTENT="Radebaugh, Jacqueline">
<META NAME="DC.Contributor" LANG="en" CONTENT="Library of Congress
  Network Development Staff">
<META NAME="DC.Contributor" LANG="en" CONTENT="MARC Standards Office
  Staff">
<META NAME="DC.Date" LANG="en" CONTENT="2004">
```

```
&lt;META NAME="DC.Type" LANG="en" CONTENT="text"&gt;
&lt;META NAME="DC.Format" LANG="en" CONTENT="application/pdf"&gt;
&lt;META NAME="DC.Identifier" LANG="en" CONTENT="ISBN:1880124629"&gt;
&lt;META NAME="DC.Source" LANG="en" CONTENT="Metadata Made Simpler: A Guide
for Libraries"&gt;

&lt;META NAME="DC.Language" LANG="en" CONTENT="en"&gt;
&lt;META NAME="DC.Relation" LANG="en" CONTENT="IsBasedOn Metadata Made
Simpler: A guide for libraries"&gt;

&lt;META NAME="DC.Rights" LANG="en" CONTENT="Copyright 2004 National
Information Standards Organization"&gt;
```

## 2. Brave New FRBR World. Patrick Le Bœuf. 2003

```
&lt;META NAME="DC.Title" LANG="en" CONTENT="Brave New FRBR World"&gt;
&lt;META NAME="DC.Creator" LANG="en" CONTENT="Le Bœuf, Patrick"&gt;
&lt;META NAME="DC.Subject" LANG="en" CONTENT="Functional Requirements for
Bibliographic Records (FRBR)"&gt;

&lt;META NAME="DC.Description" LANG="en" CONTENT="FRBR(1) is the result of a
study about the functional requirements for bibliographic records
undertaken from 1992 through 1997 by a group of experts and consultants
as a consequence of one of the 9 resolutions adopted in 1990 at the
Stockholm Seminar on Bibliographic Records."&gt;

&lt;META NAME="DC.Publisher" LANG="en" CONTENT="Die Deutsche Bibliothek"&gt;
&lt;META NAME="DC.Date" LANG="en" CONTENT="1998-05"&gt;
&lt;META NAME="DC.Type" LANG="en" CONTENT="text"&gt;
&lt;META NAME="DC.Format" LANG="en" CONTENT="application/pdf"&gt;
&lt;META NAME="DC.Identifier" LANG="en"
CONTENT="http://www.ddb.de/news/pdf/papers_leboeuf.pdf"&gt;

&lt;META NAME="DC.Language" LANG="en" CONTENT="en"&gt;
```

## 3. The Digital Object Identifier System. [Website]

```
&lt;META NAME="DC.Title" LANG="en" CONTENT="The Digital Object Identifier
System"&gt;

&lt;META NAME="DC.Creator" LANG="en" CONTENT="The International DOI
Foundation (IDF)"&gt;
```

&lt;META NAME="DC.Subject" LANG="en" CONTENT="digital object identifier system; DOI; persistent identification; intellectual content; metadata; content suppliers; electronic commerce; automated media management"&gt;

&lt;META NAME="DC.Description" LANG="en" CONTENT="The Digital Object Identifier (DOI) is a system for identifying content objects in the digital environment. DOIs are names assigned to any entity for use on digital networks. They are used to provide current information, including where they (or information about them) can be found on the Internet. Information about a digital object may change over time, including where to find it, but its DOI will not change."&gt;

&lt;META NAME="DC.Publisher" LANG="en" CONTENT="International DOI Foundation"&gt;

&lt;META NAME="DC.Date" LANG="en" CONTENT="2004-11-23"&gt;

&lt;META NAME="DC.Type" LANG="en" CONTENT="text"&gt;

&lt;META NAME="DC.Format" LANG="en" CONTENT="text/html"&gt;

&lt;META NAME="DC.Identifier" LANG="en"  
CONTENT="http://www.doi.org/index.html"&gt;

&lt;META NAME="DC.Language" LANG="en" CONTENT="en"&gt;

&lt;META NAME="DC.Rights" LANG="en" CONTENT="Copyright International DOI Foundation. All rights reserved."&gt;

#### 4. Trusted Digital Repositories: Attributes and Responsibilities. RLG-OCLC. 2002.

&lt;META NAME="DC.Title" LANG="en" CONTENT="Trusted digital Repositories: Attributes and Responsibilities"&gt;

&lt;META NAME="DC.Creator" LANG="en" CONTENT="RLG/OCLC Working Group on Digital Archive Attributes"&gt;

&lt;META NAME="DC.Subject" LANG="en" CONTENT="digital repositories; OAIS; Open Archival Information Systems; Research Libraries Group; RLG; OCLC; Online Computer Catalog Center"&gt;

&lt;META NAME="DC.Description" LANG="en" CONTENT="A trusted digital repository is one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future. In this report the working group has articulated a framework of attributes and responsibilities for trusted, reliable, sustainable digital repositories capable of handling the range of materials held by large and small research institutions."&gt;

&lt;META NAME="DC.Publisher" LANG="en" CONTENT="Research Libraries Group"&gt;

&lt;META NAME="DC.Contributor" LANG="en" CONTENT="Beagrie, Neil"&gt;

&lt;META NAME="DC.Contributor" LANG="en" CONTENT="Doerr, Dr. Marianne"&gt;

&lt;META NAME="DC.Contributor" LANG="en" CONTENT="Hedstrom, Dr. Margaret"&gt;

&lt;META NAME="DC.Contributor" LANG="en" CONTENT="Jones, Maggie"&gt;  
 &lt;META NAME="DC.Contributor" LANG="en" CONTENT="Kenney, Anne"&gt;  
 &lt;META NAME="DC.Contributor" LANG="en" CONTENT="Lupovici, Catherine"&gt;  
 &lt;META NAME="DC.Contributor" LANG="en" CONTENT="Russell, Kelly"&gt;  
 &lt;META NAME="DC.Contributor" LANG="en" CONTENT="Webb, Colin"&gt;  
 &lt;META NAME="DC.Contributor" LANG="en" CONTENT="Woodyard, Deborah"&gt;  
 &lt;META NAME="DC.Contributor" LANG="en" CONTENT="Dale, Robin"&gt;  
 &lt;META NAME="DC.Contributor" LANG="en" CONTENT="Bellinger, Meg"&gt;  
 &lt;META NAME="DC.Date" LANG="en" CONTENT="2002-05"&gt;  
 &lt;META NAME="DC.Type" LANG="en" CONTENT="text"&gt;  
 &lt;META NAME="DC.Format" LANG="en" CONTENT="application/pdf"&gt;  
 &lt;META NAME="DC.Identifier" LANG="en"  
     CONTENT="http://www.rlg.org/longterm/repositories.pdf"&gt;  
 &lt;META NAME="DC.Source" LANG="en" CONTENT="Reference Model for an Open  
     Archival Information System (OAIS)"&gt;  
 &lt;META NAME="DC.Language" LANG="en" CONTENT="en"&gt;  
 &lt;META NAME="DC.Relation" LANG="en" CONTENT="IsBasedOn Reference Model for  
     an Open Archival Information System (OAIS)"&gt;  
 &lt;META NAME="DC.Rights" LANG="en" CONTENT="Copyright 2002 Research  
     Libraries Group, All rights reserved"&gt;

## 5. An Introduction to the Resource Description Framework. Eric Miller. 1998

&lt;META NAME="DC.Title" LANG="en" CONTENT="An Introduction to the Resource  
     Description Framework"&gt;&lt;META NAME="DC.Creator" LANG="en"  
     CONTENT="Miller, Eric"&gt;  
 &lt;META NAME="DC.Subject" LANG="en" CONTENT="Resource Description  
     Framework"&gt;  
 &lt;META NAME="DC.Description" LANG="en" CONTENT="The Resource Description  
     Framework (RDF) is an infrastructure that enables the encoding,  
     exchange and reuse of structured metadata. RDF is an application of XML  
     that imposes needed structural constraints to provide unambiguous  
     methods of expressing semantics. RDF additionally provides a means for  
     publishing both human-readable and machine-processable vocabularies  
     designed to encourage the reuse and extension of metadata semantics  
     among disparate information communities. The structural constraints RDF  
     imposes to support the consistent encoding and exchange of standardized  
     metadata provides for the interchangeability of separate packages of  
     metadata defined by different resource description communities."&gt;  
 &lt;META NAME="DC.Publisher" LANG="en" CONTENT="D-Lib Magazine"&gt;  
 &lt;META NAME="DC.Contributor" LANG="en" CONTENT="Jul, Erik"&gt;  
 &lt;META NAME="DC.Contributor" LANG="en" CONTENT="Lassila, Ora"&gt;  
 &lt;META NAME="DC.Contributor" LANG="en" CONTENT="LeVan, Ralph"&gt;

```

<META NAME="DC.Contributor" LANG="en" CONTENT="Guha, R.V.">
<META NAME="DC.Contributor" LANG="en" CONTENT="Schloss, Bob">
<META NAME="DC.Contributor" LANG="en" CONTENT="Swick, Ralph">
<META NAME="DC.Contributor" LANG="en" CONTENT="Weibel, Stuart">
<META NAME="DC.Date" LANG="en" CONTENT="1998-05">
<META NAME="DC.Type" LANG="en" CONTENT="text">
<META NAME="DC.Format" LANG="en" CONTENT="text/html">
<META NAME="DC.Identifier" LANG="en" CONTENT="ISSN 1082-9873">
<META NAME="DC.Language" LANG="en" CONTENT="en">
<META NAME="DC.Rights" LANG="en" CONTENT="Copyright 1998 Eric Miller">

```

## Discussion

The growing practice of gleaning metadata from electronic resources is still evolving. Considering the differences in various metadata systems, with ERC and DC being just two of the numerous systems available, it appears obvious that there is a recognized need for metadata but a continuing disagreement in methodology. .

The simplicity of ERC is misleading. As demonstrated in the examples included in this paper, which were purposely left in the most simplistic form having just the basic story segment, minimal information is provided about the entity. Additional stories are then left up to the discretion of the various individuals developing the metadata which increases the possibility of inconsistencies. For instance, with no standardized qualifiers for elements, two individuals developing metadata for similar intellectual entities, both containing illustrations, may decide on two different label qualifiers for the same type of information. One individual may choose “erc: who/illustrator: Smith, John,” while the other may choose “erc: who/artist: Smith, John.” The metadata would then be stored in separate locations resulting in poor retrieval response when a user is searching for entities with John Smith drawings or illustrations. Additionally, providing limited standardized criteria for searches increases the chances the user will be able to retrieve the sought after information as opposed to unlimited element labels, many of which may never occur to the user.

Conversely, DC metadata is more structured and provides limited standardized criteria while including what may be considered the most critical components about an information entity. The 15 DC elements provide a well formed foundation of information about the intellectual entity by placing emphasis on fundamental information while including a couple more esoteric components, such as “Coverage” and “Rights.” Additionally, some of the elements are further characterized by the use of controlled vocabularies to promote and maintain interoperability. Controlled vocabularies help to maintain the quality of the data in the metadata records by ensuring consistent and reliable information.

The true challenge for information specialists producing the metadata is being able to discern the appropriate data for each of the elements in the different intellectual entities. While some of the information may be easily recognizable, such as “Title” or “Publisher”, some of the elements are

not quite as easily established. The element “Description” is a more subjective term and requires the developer of the metadata to have a certain level of understanding of the resource. The “Contributor” element is also not always easily determined from the resource; the key being that the contribution must be of a significant nature. Finally, the element “Coverage” seems to be a rarely used and very unusual characteristic that is found in only some resources.

Utilizing a generic metadata editor, NoteTab Light, helps the novice properly structure metadata records in either ERC or DC by providing the appropriate prompts for the required and optional elements. As with any generic version of software the capabilities are limited, but for a person with limited experience it is a good place to start.

## **Conclusion**

Lessons learned from this assignment include a better understanding of metadata and the fine nuances associated with proper extraction of information. Using two different systems for developing metadata provided the opportunity to compare structures and methods and helped determine preferences. Additionally, this assignment afforded familiarization with terminology, resources for further investigation and study, and the introduction to an exciting and evolving topic. The excitement of discovery is somewhat tempered by the suspicion that what has been covered to this point is merely the tip of the iceberg.